

Thiel, Felicitas; Cortina, Kai S.; Pant, Hans Anand

Steuerung im Bildungssystem im internationalen Vergleich

Fatke, Reinhard [Hrsg.]; Oelkers, Jürgen [Hrsg.]: *Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart*. Weinheim; u.a. : Beltz Juventa 2014, S. 123-138. - (Zeitschrift für Pädagogik, Beiheft; 60)



Quellenangabe/ Reference:

Thiel, Felicitas; Cortina, Kai S.; Pant, Hans Anand: Steuerung im Bildungssystem im internationalen Vergleich - In: Fatke, Reinhard [Hrsg.]; Oelkers, Jürgen [Hrsg.]: *Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart*. Weinheim; u.a. : Beltz Juventa 2014, S. 123-138 - URN: urn:nbn:de:0111-opus-90914 - DOI: 10.25656/01:9091

<https://nbn-resolving.org/urn:nbn:de:0111-opus-90914>

<https://doi.org/10.25656/01:9091>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipt.de
Internet: www.pedocs.de

ZEITSCHRIFT FÜR PÄDAGOGIK

**Das Selbstverständnis
der Erziehungswissenschaft:
Geschichte und Gegenwart**

Zeitschrift für Pädagogik · 60. Beiheft

Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart

Herausgegeben von
Reinhard Fatke und Jürgen Oelkers

BELTZ JUVENTA

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, bleiben dem Beltz-Verlag vorbehalten.

Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genutzte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, bei der die einzelnen Zahlungsmodalitäten zu erfragen sind.

© 2014 Beltz Juventa · Weinheim und Basel

www.beltz.de · www.juventa.de

Herstellung: Lore Amann

Satz: text plus form, Dresden

Druck und Bindung: Beltz Bad Langensalza GmbH, Bad Langensalza

Printed in Germany

ISSN 0514-2717

Bestell-Nr. 443502

Inhaltsverzeichnis

Reinhard Fatke/Jürgen Oelkers

Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. Einleitung zum Beiheft	7
---	---

Wandel des Selbstverständnisses

Klaus-Peter Horn

Pädagogik/Erziehungswissenschaft der Gegenwart – Zur Entwicklung der deutschen Erziehungswissenschaft im Spiegel ihrer disziplinären Selbstreflexion (1910–2010)	14
--	----

Patrick Bühler

Einführungen in die Pädagogik. Die Entwicklung einer Gattung	33
--	----

Michael Geiss

Die angemessene Form: Abstraktion und die Suche nach einem pädagogischen Kalkül	47
--	----

Markus Rieger-Ladich

Pädagogik als kritische Theorie? Intellektuelle Stellungskämpfe nach 1945	66
--	----

Wissenschaftspraxis und Empirie

Jürgen Oelkers

Praxis und Wissenschaft. Überlegungen zur Forschungsstruktur der Erziehungswissenschaft	85
--	----

Olaf Köller

Entwicklung und Erträge der jüngeren empirischen Bildungsforschung	102
--	-----

Felicitas Thiel/Kai S. Cortina/Hans Anand Pant

Steuerung im Bildungssystem im internationalen Vergleich	123
--	-----

Heinz-Elmar Tenorth

Politikberatung und Wandel der Expertenrolle oder: Die Expertise der Erziehungswissenschaft	139
--	-----

Die Erziehungswissenschaft und ihre Nachbarwissenschaften

Bernhard Schmidt-Hertha/Rudolf Tippelt

Erziehungswissenschaft und das Verhältnis zu ihren Bezugsdisziplinen vor dem Hintergrund der Stellen- und Besetzungspraxis (1995 bis 2012)	172
--	-----

Peter Kauder

Themenkonjunkturen im Spiegel erziehungswissenschaftlicher Dissertationen und Habilitationen 1945–2009	184
---	-----

Nicole Becker

Mehr verstehen, besser handeln? Zum Verhältnis von Pädagogik und Neurowissenschaften	208
---	-----

Essay

Roland Reichenbach

Schulkritik. Eine „metaphorologische“ Betrachtung	226
---	-----

Felicitas Thiel/Kai S. Cortina/Hans Anand Pant

Steuerung im Bildungssystem im internationalen Vergleich

1. Neue Steuerungsmodelle und Steuerungsinstrumente im Bildungssystem

Seit mehr als drei Jahrzehnten wird in den meisten Industrienationen eine intensive Debatte um Effektivitäts- und Effizienzprobleme öffentlicher Verwaltungen sowie um Reibungsverluste bei der Umsetzung politischer Reformen geführt. Als Antwort auf die diagnostizierten Defizite wurden einerseits im Zuge der Implementation von *New Public Management* Entscheidungszuständigkeiten zwischen nationaler, föderaler und lokaler Ebene neu geordnet sowie eine stärkere Orientierung administrativer Prozesse an den Interessen von Anspruchsberechtigten oder Kunden ausgerichtet (Brückner & Tarazona, 2010). Andererseits verpflichteten sich einige Staaten auf ein Modell evidenzbasierter Steuerung – besonders wirkungsvoll im Jahr 1999 die von Tony Blair geführte britische Regierung unter dem Titel „Modernising Government“ (Cabinet Office, 1999) –, das die Entscheidung für bestimmte politische Maßnahmen an Wirksamkeitsnachweise knüpfte und damit nicht zuletzt eingeschliffene politische Entscheidungsrou-tinen in Frage stellte.

Zahlreiche Staaten haben in den vergangenen zwanzig Jahren Elemente des *New Public Management* sowie des Modells evidenzbasierter Steuerung im Rahmen neuer bildungspolitischer Steuerungskonzepte implementiert. Auf den Begriff gebracht, verbinden diese neuen Steuerungskonzepte Maßnahmen zur Erhöhung von Rechenschaftspflicht (*Accountability*) mit Maßnahmen zur Ausweitung von *Autonomie* auf den der Politik nachgelagerten Entscheidungsebenen. Schulen fungieren infolge einer stärkeren Dezentralisierung operativer Entscheidungszuständigkeiten als eigenständige Steuerungsakteure (Thiel, 2008a). Neue Akteure wie Institute zur Entwicklung, Durchführung und Auswertung zentraler Leistungstests oder Examen, Schulinspektoren und Schulberater treten auf den Plan, während traditionelle Akteure wie die Schulaufsicht sich mit einer Veränderung ihres Aufgabentableaus sowie der Neuvermessung ihrer Zuständigkeiten konfrontiert sehen (Altrichter & Maag Merki, 2010; McDonnell, 2013). Im historischen Rückblick lassen sich die neuen Steuerungskonzepte aufgrund der Reduzierung der direkten staatlichen Einflussnahme auf schulische und unterrichtliche Prozesse durchaus als eine Ausweitung der Spielräume für professionelle Selbstregulierung interpretieren (Thiel, 2008b), allerdings im Rahmen von Standards, die in einem von der staatlichen Politik kontrollierten und öffentlich kontrovers diskutierten Prozess definiert werden.

Weil im Vergleich zum alten Steuerungsmodell neben dem Zugang zu Bildungsressourcen die Qualität der Lerngelegenheiten sowie die erreichten Lernergebnisse er-

höhte Aufmerksamkeit erfahren, sind eine Verständigung über zu erreichende Standards sowie eine valide Messung von *Outcomes* zentrale Voraussetzungen des neuen Steuerungsmodells. Gemeinsam ist den verschiedenen nationalen Steuerungsstrategien folglich die Einführung von Bildungsstandards und standardbasierten Tests (vgl. Altrichter & Kanape-Willingshofer, 2012). Verschiedene nationale Strategien unterscheiden sich allerdings erheblich hinsichtlich ihrer Ausrichtung. In sogenannten *High-Stakes*-Systemen erfolgt eine testbasierte Zurechnung von Effekten auf Einzelschulen und Lehrkräfte mit dem Ziel der Akkreditierung von Schulen, der Entscheidung über Mittelzuweisung, über Einstellung und Bezahlung von Lehrkräften sowie – im Fall fortgesetzt auftretender Defizite – über die Entlassung von Lehrkräften und einen Austausch des Managements. In sogenannten *Low-Stakes*-Systemen haben die Testdaten dagegen die Funktion eines Feedbacks für professionelle Akteure, auf dessen Grundlage Prozesse der Schul- und Unterrichtsentwicklung angestrengt werden sollen. Während *High-Stakes*-Systeme vor allem auf die Steuerungswirkung von Kontroll- oder Marktmechanismen setzen, betonen *Low-Stakes*-Systeme die Selbstregulierungskraft der professionellen Gemeinschaft. Van Ackeren (2003) unterscheidet zwischen dem *Pressure*-Ansatz, der auf öffentliche Rechenschaftslegung, freie Schulwahl und Wettbewerb zwischen Schulen setzt, und dem *Support*-Ansatz, der auf vertrauliche Rückmeldung setzt und damit die Hoffnung eines Anstoßes schulischer und unterrichtlicher Entwicklung verbindet.

Entgegen manchen Behauptungen geht die Einführung von standardbasierten Steuerungsmodellen nicht grundsätzlich mit einer Reduktion bürokratischer Regulierung einher (vgl. Vincenti & Geiss, 2012). Sowohl in *High-Stakes*- als auch in *Low-Stakes*-Systemen spielt die staatliche Bildungsverwaltung nach wie vor eine große Rolle, nicht allein aufgrund der Kontrolle über den Prozess der Standardfestlegung, der Akkreditierung von Unterrichtsmaterialien und Fortbildungen, sondern auch aufgrund der Überwachung der vielfältigen Dokumentations- und Berichtspflichten von Schulen. Allerdings ist kaum zu bestreiten, dass sich Funktion und Rolle staatlicher Schulaufsicht im System der Neuen Steuerung stark verändert haben – etwa wenn sie in *low-performing schools* direkt interveniert (Wong, 2013) – und dass neue Akteure wie Schulinspektorate bürokratische Regulierungsfunktionen übernehmen.

Die neue Steuerungsphilosophie speist sich aus unterschiedlichen theoretischen Ansätzen. Prominent sind vor allem die Bezugnahme auf die Neue Institutionenökonomik, genauer die *Agenturtheorie*, auf der einen und die Bezugnahme auf die *Professionstheorie* auf der anderen Seite.

Agenturtheoretische Ansätze argumentieren, dass in einer Prinzipal-Agent-Beziehung grundsätzlich ein Opportunismusrisiko angelegt ist, weil die Interessen zwischen einem Prinzipal (Auftraggeber) und einem Agenten (Beauftragter) divergieren können und die Informationen des Prinzipals über die Leistungserbringung des Agenten grundsätzlich beschränkt sind (Ebers & Gotsch, 2006; Wößmann, Luedemann, Schuetz & West, 2009). Prinzipal-Agenten-Beziehungen können z. B. zwischen Schulaufsicht und Schulleitung, zwischen Schulleitung und Lehrkräften oder zwischen Eltern und Schulen bestehen. Dem theoretischen Ansatz zufolge erhöhen zuverlässige Daten die Transparenz für den Prinzipal hinsichtlich der Leistungserbringung des Agenten. Auf der Grund-

lage aussagekräftiger Informationen über die (potenziellen) Leistungen des Agenten können einerseits kompetente Agenten rekrutiert und andererseits Anreize und Sanktionen mit dem Ziel eingesetzt werden, die Interessen zwischen Prinzipal und Agenten in Bezug auf die Leistungserbringung zur Deckung zu bringen. Die Herausforderungen für die Implementation dieses Steuerungsansatzes im Schulsystem liegen auf der Hand. Zum einen muss sichergestellt werden, dass aussagekräftige Informationen über die Leistung der Lehrkräfte verfügbar sind. Zum anderen müssen die möglichen Effekte dieses Steuerungsmodells mit den Agenturkosten, also dem finanziellen, personellen und sozialen Aufwand der Informationsbeschaffung, abgeglichen werden. Um die Aussagekraft der Informationen zu gewährleisten, müssen Daten kontextualisiert werden. In den meisten Fällen wird die Schulperformanz über aggregierte Leistungswerte der Schülerinnen und Schüler operationalisiert. Die Informationen über Leistungsoutputs werden dabei um Informationen zur Eingangszusammensetzung der Schülerschaft sowie zur Prozessqualität auf der Ebene der Schule und des Unterrichts ergänzt, die von Schulinspektionen und Schulleitungen erhoben werden. Interessanterweise existieren bislang nur Vermutungen dazu, wie die Bilanz eines Abgleichs von Aufwand und Effekten der Evaluationsverfahren ausfällt.

In einigen Staaten erfolgt eine Veröffentlichung der Schulleistungsdaten mit dem Argument, dass sich Transparenz nicht nur auf die Reduktion der Informationsasymmetrie zwischen Lehrkräften und Schulaufsicht bzw. Schulleitung beschränken dürfe, sondern auch Eltern einbeziehen müsse, um die Entscheidungsmacht von Anspruchsberechtigten staatlicher Leistungen zu stärken (McDonnell, 2013). Unter den Bedingungen der freien Schulwahl wird durch die Veröffentlichung der Daten eine Wettbewerbssituation erzeugt bzw. verschärft. Eine Voraussetzung dafür, dass diese Wettbewerbssituation im Sinne der Qualitätsverbesserung pädagogischer Leistung wirksam wird, sind allerdings Entscheidungsspielräume auf der Schulebene, die es den Schulen als Anbietern pädagogischer Leistung erlauben, flexibel auf Nachfrageimpulse zu reagieren. Im Innenverhältnis bilden Schulen – Entscheidungsspielräume vorausgesetzt – der Logik der Marktanpassung zufolge entsprechende Strukturen aus, die eine effiziente Anpassung an Marktdynamiken erlauben (Thiel, 2008a, 2008b; vgl. auch Belfield & Levin, 2009). Entscheidungsspielräume auf der Schulebene sind in der Logik der Prinzipal-Agenten-Theorie besonders dort relevant, wo Schulen aufgrund einer besseren Kenntnis von Bedarfen und erforderlichen Kompetenzen einen Wissensvorsprung haben. Das ist insbesondere im Bereich der Personalrekrutierung und -entwicklung der Fall (Wößmann et al., 2009).

Gegen die Schaffung von Wettbewerbsbedingungen im Schulsystem wird vielfach eingewandt, dass einerseits die Voraussetzungen zur Rezeption von veröffentlichten Leistungsdaten bei den Nachfragern nach staatlichen Bildungsangeboten nicht gleich verteilt sind und dass andererseits die sozialräumliche Segregation der Wohnbevölkerung der freien Wahl der Schule Grenzen setzt. Die Schaffung von Marktbedingungen gehe aus diesem Grund zwangsläufig mit der Entstehung von „Restschulsituationen“ und damit einer Verschärfung von Ungleichheit einher (Heinrich, Altrichter & Sokoup-Altrichter, 2011). Von Kritikern des *High-Stakes*-Steuerungsansatzes werden

häufig auch die Zuverlässigkeit und die Validität der Leistungsdaten hinsichtlich der Beurteilung der professionellen Leistung in Zweifel gezogen. Die Messgenauigkeit von Effekten, die Lehrkräften oder Schulen zurechenbar sind (*added values*; vgl. Amrein-Beardsley, 2008), sei grundsätzlich als problematisch anzusehen, da die *Value-Added*-Indikatoren im Zeitverlauf statistisch-methodisch bedingte Schwankungen aufwiesen, die sie als Kennziffer für Lohnanteils- oder andere Mittelzuweisungen fraglich erscheinen lassen (Raudenbush, 2004). Hinsichtlich der Validität werde mit standardbasierten Tests zum einen nur ein kleiner Ausschnitt des Spektrums professioneller Leistungen von Lehrkräften erfasst; zum anderen bestünden erhebliche Bedenken, weil unerwünschte Einflussnahmen wie *Teaching-to-the-Test* oder gar Manipulation und Betrug nie auszuschließen seien und durch *High-Stakes*-Evaluationen sogar verstärkt würden (Amrein-Beardsley, Berliner & Rideau, 2010; Fullan, 2009; Koretz, 2008, 2011).

Der andere theoretische Bezugsrahmen für Programme der Neuen Steuerung im Bildungssystem ist die *Professionstheorie* (vgl. Schimank, 2014; Tenorth, 1989; Terhart, 1992). Auch in *High-Stakes*-Systemen wie etwa in den USA finden sich Elemente des Ansatzes der professionellen Lerngemeinschaft (Leithwood & Menzies, 1998), die als klassischer Mechanismus der Qualitätssicherung in professionellen Organisationen gilt (Freidson, 2001). Professionelle Selbstregulierung funktioniert – anders als bürokratische Kontrolle oder Wettbewerb – auf der Basis von allgemeinen, durch die Profession definierten und kontrollierten professionellen Standards. Voraussetzung für professionelle Selbstregulierung sind (a) ein Kernbestand professionellen Wissens und professioneller Normen, der die Qualität der professionellen Leistung begründet, (b) ein professionelles Ethos, das die Leistungserbringung im Sinne dieser Normen motiviert, und (c) Institutionen professioneller Selbstkontrolle, die grobe Abweichungen von den Normen feststellen und Missbrauch professioneller Autonomie sanktionieren können.

Auf der Grundlage von Informationen über den Bedarf von Klienten wird professionelles Wissen adaptiert, um die professionelle Leistung bedarfsgerecht zuzuschneiden. Diese Adaptation geschieht im Horizont eines spezifischen Falls (eines Schülers bzw. einer Schülerin, einer Klasse oder einer Schule). Das Modell der professionellen Selbstregulierung wird durch die klassischen Professionen repräsentiert, insbesondere durch die ärztliche Profession (Freidson, 2001).

Leistungsdaten sind dem Ansatz der professionellen Selbstregulierung zufolge dann wertvolle Informationen, wenn die Profession die zentrale Rolle bei der Definition der Standards spielt und die Datenhoheit allein bei der Profession liegt. In der Logik des Modells der professionellen Selbstregulierung sorgen nicht ein gut funktionierendes schulisches Management oder externe Marktmechanismen für die Qualität der Leistungserbringung, sondern eine gemeinschaftliche Verständigung über die Ziele schulischer Arbeit, die *Peer-Evaluation* oder interne Evaluation sowie die fallbezogene Entwicklung passgenauer Interventionen. Schulprogramm, schuleigenes Curriculum und interne Evaluation zählen deshalb ebenso wie kollaborative Unterrichtsentwicklung zum Kernbestand der Instrumente professioneller Selbstregulierung (vgl. Thiel, 2008a; Thiel & Thillmann, 2012). Externe Kontrolle kann aus der Sicht dieses Ansatzes Qualitätsentwicklung unterminieren, weil sie die professionelle Autorität und damit Moti-

vation infrage stellt und so die Bindungswirkung des professionellen Ethos außer Kraft setzt (vgl. Hargreaves, 1995).

Eine Übertragung des Modells der ärztlichen auf die Lehrerverberfession ist allerdings aus mehreren Gründen problematisch: erstens, weil über den Kernbestand professionellen Wissens und über professionelle Standards kein uneingeschränkter Konsens zwischen den Mitgliedern der Profession besteht; zweitens, weil kein institutionalisiertes professionelles Ethos (Professionskammer; vgl. Schlömerkemper, 2009) existiert; drittens, weil Mechanismen der Wissensapplikation (z. B. evidenzbasierte Entscheidungen und Konsultation) und der Wissensentwicklung (z. B. evidenzbasierte Fort- und Weiterbildung) nicht oder nur ansatzweise etabliert sind. Auf der anderen Seite sind die Spielräume für professionelle Selbstregulierung bei pädagogischen Professionen durch staatliche Kontrolle deutlich beschränkt. Die Lehrerverberfession wird deshalb aus gutem Grund als *heteronome professionelle Organisation* bezeichnet (Waters, 1989).

Auch der Ansatz der professionellen Selbstregulierung weist systematische Schwächen auf: Weil gemeinschaftliche Regulierung auf „emotionale[r] Interdependenz“ (Wiesenthal, 2000, S. 55) beruht, sind Wertbindung und Vertrauen zentrale gemeinschaftliche Ressourcen. Die grundsätzliche Schwäche gemeinschaftlicher Regulation besteht deshalb – neben einer gewissen „kognitive[n] Schließung“ (Wiesenthal, 2000, S. 60) – in „informellen wechselseitigen Nichtangriffspakten“ (Schimank, 1994, S. 22), die Schutz vor der Beschädigung von Autonomie und Vertrauen bieten, was allerdings häufig dazu führt, dass individuelle Schwächen in der Leistungserbringung nicht, z. B. im Rahmen systematischer kollegialer Hospitationen, aufgedeckt werden (vgl. Thiel, 2008b). Empirische Befunde aus Schweden, wo bildungspolitische Entscheidungen weitgehend dezentralisiert wurden, verweisen zudem darauf, dass die Kooperation im Kollegium bisher nicht im Sinne des Modells gemeinsamer Planung, Evaluation und Entwicklung funktioniert hat (vgl. Kotthoff, 2003, S. 327).

2. Steuerungsmodelle im Vergleich

Obwohl sich die Reformakteure in allen OECD-Staaten auf die Modelle des *New Public Management* und der *evidenzbasierten Steuerung* berufen, entstehen durch die selektive Implementation und Ausgestaltung unterschiedlicher Instrumente vor dem Hintergrund der jeweiligen national-kulturellen Traditionen der Schulpolitik ganz unterschiedliche Konfigurationen. Diese Steuerungskonfigurationen können danach beschrieben werden, wie sie die unterschiedlichen Regulierungsmechanismen (Wettbewerb, Bürokratie und professionelle Gemeinschaft) kombinieren (Thiel, 2008a). Schmid, Hafner und Pirolt (2007) haben eine Typologie von Steuerungsregimen vorgelegt, die verdeutlicht, wie in verschiedenen OECD-Staaten die Steuerungsprinzipien Weisung, Wettbewerb und Partizipation ins Verhältnis gesetzt werden. Sie unterscheiden dabei den *Bürokratiety*, den *Effizienzty* und den *Legitimitätsty*. Während in Deutschland und Österreich nach wie vor die bürokratische Regulierung dominiert, allerdings kombiniert mit Ansätzen des *School Empowerment*, setzen andere Staaten wie Finnland oder Polen verstärkt auf

Partizipation im Sinne von *Local Empowerment* (Polen) und *Local Empowerment* in Kombination mit *School Empowerment* (Finnland). Beide Staaten werden dem *Legitimitätstyp* bildungspolitischer Steuerung zugerechnet. Allein Neuseeland, das den *Effizienztyp* repräsentiert, setzt bei der Steuerung des Schulsystems vorwiegend auf Wettbewerbsmechanismen (Schmid et al., 2007). England, das von den Autoren ebenfalls dem Effizientyp zugerechnet wird, kombiniert staatliche Kontrolle (*strong state*) mit Marktmechanismen wie freier Schulwahl (*free market*) (Kotthoff, 2003, S. 188). Laut Kotthoff zeichnet sich das englische System durch eine alle Bereiche des Bildungssystems umgreifende Evaluationsbürokratie aus, die im Unterschied zum US-amerikanischen Modell neben standardbasierten Testverfahren auch auf eine umfängliche staatlich kontrollierte Evaluation von Input- und Prozessmerkmalen durch Schulinspektionen setzt (Kotthoff, 2003, S. 189).

Zu einer etwas anderen, aber prinzipiell an Schmid et al. (2007) anschlussfähigen Gruppierung kommen Windzio, Sackmann und Martens (2005), die den Aspekt der Bildungsfinanzierung stärker berücksichtigen. Sie unterscheiden sechs *Cluster der Educational Governance*: *Scandinavian Governance*, *Private Governance*, *State-based Governance*, *South European Governance*, *Anglo-Saxon-Eastern Governance* und *East-Asian Governance*. Deutschland und auch Österreich zählen mit anderen zum Cluster der *State-based Governance*. England und Neuseeland repräsentieren mit anderen Staaten das *Anglo-Saxon-Eastern Cluster* und die Niederlande sind dem *Private Governance*-Typus zugeordnet.

Eine Studie zur Systemeffizienz des Schulwesens, die 28 OECD-Staaten hinsichtlich unterschiedlicher Effizienzkriterien in den Dimensionen Ressourcenallokation, Budgetmanagement und Markteffizienz vergleicht, kommt zum Ergebnis, dass Großbritannien, Australien, Norwegen, Dänemark und die Niederlande im Vergleich zum OECD-Durchschnitt einen signifikant höheren Grad der Markteffizienz sowie der Ressourceneffizienz oder Budgeteffizienz aufweisen.

Die USA, Neuseeland, Island und Portugal zählen zu der Gruppe, die zwar hinsichtlich der Ausgabeneffizienz insgesamt über dem OECD-Durchschnitt liegt, aber nur in einer der drei Dimensionen überdurchschnittliche Effizienz bzw. in zwei Dimensionen überdurchschnittliche und gleichzeitig in einer Dimension unterdurchschnittliche Effizienz aufweist. Deutschland wird der Gruppe von Staaten zugerechnet, die in keiner der Kategorien über dem Durchschnitt liegen bzw. in einer Kategorie überdurchschnittlich effizient sind, was aber durch unterdurchschnittliche Effizienz in einer anderen Kategorie relativiert wird (Gonand, Joumard & Price, 2007).

Was die Frage der Effektivität unterschiedlicher Steuerungsinstrumente betrifft, geben Untersuchungen von Wößmann et al. (2009) im Einklang mit anderen Studien (z. B. Hanushek & Raymond, 2005) einige Hinweise auf positive Effekte zentraler Abschluss-examen und standardbasierter Tests. Die Leistungen von Schülerinnen und Schülern in Ländern, die zentrale Abschluss-examen durchführen, sind demnach signifikant besser als in Ländern ohne zentrale Abschlussprüfungen (Bishop, 1997, 1999; Fuchs & Wößmann, 2007). Dieser Effekt wurde ansatzweise auch für die deutschen Bundesländer gefunden (Jürges, Schneider & Büchel, 2005). Ähnliche Belege für die Effektivität

tät neuer Steuerungsinstrumente werden von Wößmann et al. (2009) für standardisierte Tests, für Schulinspektionen sowie für ein Monitoring der Lehrkräfte durch Schulleitungen angeführt.

Auch was die Frage der Effekte höherer Autonomiespielräume für Schulen betrifft, unterstützen die wenigen vorliegenden Befunde anscheinend die Hypothese, dass eine Dezentralisierung von Entscheidungsmacht positive Effekte auf Schülerleistungen hat (Wößmann, 2003; Fuchs & Wößmann, 2007), zumindest in entwickelten Staaten (Hanushek, Link & Wößmann, 2013). Weiterhin wurde gezeigt, dass Schülerinnen und Schüler hinsichtlich ihrer Lernleistungen vor allem in Systemen profitieren, die die Erweiterung schulischer Autonomiespielräume mit Maßnahmen der *Accountability* kombinieren (Wößmann et al., 2009). Einschränkend zu diesen bildungsökonomischen Sekundäranalysen muss jedoch angemerkt werden, dass sie in der Regel auf relativ grobe Indikatoren angewiesen sind und distale Kausalhypothesen nicht theoretisch schlüssig ableiten. Hinzu kommt, dass mehrere Studien aus dem ökonomischen Bereich die zentralen Outcome-Merkmale (Schülerleistung) nur in aggregierter Form und daher nicht individualbezogen und mehrebenenanalytisch modellieren (können), sodass sie sensibel für ökologische Fehlschlüsse sind (vgl. Stanat & Lüdtke, 2008).

Nicht nur in der agenturtheoretischen, sondern auch in der professionstheoretischen Perspektive werden Schulen als zentrale Steuerungsakteure im Schulsystem betrachtet. Jenseits der allgemeinen Befunde zu Effekten von Schulautonomie ist deshalb eine differenziertere Beschreibung der tatsächlichen Autonomiespielräume in den unterschiedlichen Ländern von hohem Interesse. Eine neuere Studie von Brauckmann (2012) zeigt, dass zwischen Staaten große Unterschiede in der Dezentralisierung von Entscheidungen bestehen. Ob diese Verlagerung im Sinne einer *Devolution*, also einer Übertragung von Entscheidungsvollmachten im Sinne nicht-zustimmungspflichtiger Entscheidungen, oder der *Delegation*, also einer Dezentralisierung im Sinne einer an Vorgaben gebundenen „geliehenen“ Entscheidungsmacht, erfolgt, macht hinsichtlich der Signatur des Steuerungsregimes einen großen Unterschied. Brauckmann kann zeigen, dass nur in vier von sieben untersuchten europäischen Ländern weitreichende Entscheidungen hinsichtlich der Unterrichtsorganisation, der Personalentwicklung und der Ressourcen auf der Schulebene getroffen werden. Nur in England und den Niederlanden bestehen auf Schulebene in allen Bereichen echte Entscheidungsspielräume. Dagegen weisen Deutschland, Norwegen und Italien eher zentralisierte Entscheidungsstrukturen auf, Schulen wird hier nur die „Umsetzungsverantwortung“ (Brauckmann, 2012, S. 88) überlassen. Was Personalangelegenheiten betrifft, werden laut der durchgeführten Expertenbefragung in Deutschland überhaupt keine Entscheidungen autonom auf Schulebene getroffen. Von Gestaltungsautonomie, ob im Sinne eines marktwirtschaftlichen Steuerungsmodells oder eines Modells professioneller Selbststeuerung, kann hier demnach keine Rede sein, da die entscheidenden Gestaltungsmittel nicht zur Verfügung stehen.

Insgesamt lassen sich aus den derzeit verfügbaren Forschungsergebnissen keine validen, eindeutigen Aussagen ableiten, in welcher Austarierung *Pressure-* bzw. *Support-*Elemente bei der Steuerung von Bildungssystemen unter der Zielperspektive der Verbesserung von Leistungsergebnissen von Schülerinnen und Schülern effektiv sind. Im

nächsten Abschnitt sollen durch Kontrastierung am Fallbeispiel USA die Potenziale und Risiken der aktuellen Steuerungspraxis im deutschen Bildungssystem näher charakterisiert werden.

3. Reformen mittlerer Reichweite: zwischen *Pressure* und *Support*

So hilfreich es auch ist, die Szenarien des mittelfristigen Wandels durch die tiefgreifenden Reformen im deutschen Bildungswesen entlang der Pole *High-* und *Low-Stakes* zu ordnen oder der Dichotomie von *Pressure* und *Support*, so birgt diese Betrachtung auch die Gefahr, dass man neugeschaffene Strukturdefizite konzeptionell nicht hinreichend erfasst. In diesem Abschnitt wird die These untersucht, dass sich möglicherweise gerade das Scharniersystem von *Large-Scale Assessments* und an Bildungsstandards orientierten, schulbezogenen Vergleichsarbeiten in Deutschland als Reformhemmnis erweisen könnte.

Wir haben oben dargestellt, dass der *High-Stakes*-Steuerungsansatz innerhalb des Bildungsmonitorings darauf beruht, dass mit aussagekräftigen Daten die Effektivität der Schulen bzw. von Lehrpersonen transparent gemacht und mit Anreiz- und Sanktionsmitteln kombiniert werden kann. Transparenz hat in diesem Zusammenhang zwei Bedeutungen, nämlich Öffentlichkeit und Vergleichbarkeit. Für das rationale Handeln der Schulverwaltung sind beide Aspekte von nicht zu unterschätzender Bedeutung, weil sie verfahrenslegitimierend genutzt werden können. Im Sinne der Systemtheorie von Niklas Luhmann (1983) müssen herrschaftswirksame Verfahren der Entscheidungsfindung innerhalb öffentlicher Institutionen nicht „objektiv richtig“ oder „wissenschaftlich begründbar“ sein, sondern müssen sich in demokratischen Gesellschaften lediglich hinreichend legitimieren. In politisch-administrativen Subsystemen wie der Schulverwaltung wird Verwaltungshandeln überwiegend durch Verfahren festgelegt, die im politischen Diskursprozess entwickelt wurden und grundsätzlich revidierbar sind. Die im doppelten Sinne transparenten *High-Stakes*-Informationen geben der Schuladministration einen erstaunlich breit legitimierten Handlungsspielraum: Eine Schule mit überdurchschnittlich guten Testleistungen könnte belohnend alimentiert werden, oder man könnte ihr im Gegenteil auch Ressourcen entziehen, weil andere Schulen die Unterstützung dringender brauchen. Letztgenanntes dürfte aufgrund der öffentlichen Transparenz politisch riskant sein, weil das meritokratische Prinzip in Deutschland (wie in anderen westlichen Ländern) den öffentlichen Diskurs dominiert und demzufolge Leistung (hier: überdurchschnittliche Testleistung) belohnt werden soll. Die als schwach gekennzeichneten Schulen können entweder mit Sanktionen bedroht werden oder, verbunden mit der Erwartung von zügig sichtbaren Verbesserungen, zusätzliche Mittel erhalten.

Wohl keine andere Gesetzgebung hat weltweit die Schwächen dieser Legitimation durch Verfahren besser aufgezeigt als die in der Praxis gescheiterte Novelle des *Elementary and Secondary Education Act* (ESEA) in den USA, besser bekannt als *No Child Left Behind Act* von 2001 (NCLB; vgl. U. S. Department of Education, 2002). Ohne direkt nachvollziehbaren Bezug auf empirische Forschung zu den Inhalten und Zeitho-

rizonten von Schulreformen sah NCLB ein vierstufiges Sanktionsmodell für Schulen vor, die nicht eine „adäquate jährliche Verbesserung“ (*adequate yearly progress, AYP*) auf der Ebene der Schülerleistungen nachweisen konnten. AYP sollte von den jeweiligen Landesregierungen so festgelegt werden, dass nach zwölf Jahren Schulentwicklung alle Schülerinnen und Schüler eines Jahrgangs die jahrgangsweise definierten Mindeststandards in den Kernfächern erreichen. An die Erreichbarkeit dieses Ziels bis 2014 hat wohl ernsthaft keiner der politisch Verantwortlichen geglaubt. Doch ganz im Sinne N. Luhmanns hatte NCLB einen enormen Machtzuwachs der Schulverwaltung gegenüber den Schulen zur Folge, weil die Gesetzgebung drakonische Maßnahmen wie komplettes Ersetzen der Lehrerschaft seitens der Verwaltung legitimiert, und dies formal bereits nach vier Jahren des Nichterreichens von AYP. Fairerweise muss man darauf hinweisen, dass die ersten beiden Stufen des Sanktionsmodells in NCLB eher dem *Support*-Gedanken verpflichtet sind, weil Schulen nach zwei Jahren ohne AYP (*In-Need-of-Improvement*-Status) zunächst eine Art Schulentwicklungsplan für jedes Fach vorlegen müssen, in dem das Verbesserungssoll nicht erreicht wurde. Doch diese Art von Unterstützungsangeboten ist im Sinne van Ackerens (2003) eher dem *Pressure*-Ansatz zuzuordnen als dem *Support*-Ansatz, weil es oktroyierte Maßnahmen sind, für die zudem keine föderalen Haushaltsmittel bereitgestellt wurden und die somit in aller Regel durch die Schule selbst zu finanzieren waren. Trotz maximaler doppelter Transparenz durch Öffentlichkeit/Vergleichbarkeit ist die erhoffte Breitenwirkung unbeschadet einiger regionaler Anfangserfolge ausgeblieben (vgl. Mintrop & Sunderman, 2012). Die Schulverwaltungen haben mit der Zeit ohne erkennbaren Widerstand der Regierung in Washington davon abgesehen, den Sanktionsspielraum auszuschöpfen, um ein „Schulchaos“ zu vermeiden; ein solches wäre unweigerlich die Folge gewesen, weil in Krisenregionen wie Detroit oder Chicago jede zweite Schule von der unmittelbaren Schließung bedroht wäre.

Dem NCLB-Programm fehlte eine konkrete Vorstellung vom Arsenal möglicher Reforminitiativen auf Schulebene, der realistischen Einschätzung des Zeithorizonts ihrer Umsetzung unter den lokalen Gegebenheiten der jeweiligen Schulsituation und der entsprechenden Evaluationskriterien, die über jährliche Testungen in einigen wenigen Schulfächern hinausgingen.

Es ist durchaus sinnvoll, sich in diesem Zusammenhang die Befunde von Meta-Analysen zu vergegenwärtigen, die nicht enzyklopädisch angelegt sind, d. h. Effektstärken aller möglichen Aspekte von Schule und Unterricht vergleichen wie die jüngst vieldiskutierte Studie von Hattie (2009; vgl. Pant, im Druck), sondern gezielt die langfristigen Auswirkungen von umfassenden Reformpaketen auswerten, die in den USA als *Comprehensive School Reform (CSR)* zusammengefasst werden. Die bekanntesten Programme sind *Accelerated Schools*, *Career Academies*, *Direct Instruction*, *School Development Program* und *Success for All*. Diesen Programmen ist gemeinsam, dass eine Schule sich nicht einzelne Reformelemente herausuchen kann, sondern das gesamte Programm einführen muss, welches von Änderungen des Stundenplans über Lehrerfortbildung und -coaching bis hin zu Umstrukturierungen von Kommunikationsprozessen reicht. Borman, Hewes, Overman und Brown (2003) haben in ihrer Meta-Analyse

gezeigt, dass man selbst für solche vergleichsweise durchgreifenden Schulreformaßnahmen einen langen Atem haben muss. Der nachhaltige Erfolg auf die durch standardisierte Tests gemessene Schülerleistung stellt sich erst nach fünf bis acht Jahren ein. Die Einführung eines CSR-Programms war mit NCLB nicht kompatibel, und Schulbezirke waren genötigt, diese Schulen vom Evaluationsregime des AYP auszunehmen; dies wiederum war insofern riskant, als dem *School District* daraufhin formal alle föderalen Mittel entzogen werden konnten.

Geht man davon aus, dass CSR ein sinnvolles Instrument der Schulreform ist, so frisst hier die NCLB-Reform ihre eigenen Kinder: CSR-Reformen hätten vor einer Gesetzgebung geschützt werden müssen, die solche Reformen eigentlich stimulieren sollte. NCLB legitimiert Verwaltungshandeln durch festgelegte Verfahren, die nur auf den ersten Blick einer nachvollziehbaren Rationalität folgen. Man beachte, dass CSR-Maßnahmen sich nicht leicht in die Dichotomie von *Support* und *Pressure* nach van Ackeren (2003) einfügen, weil ihre Einführung einerseits vielfach auf Druck durch die Schulverwaltung zustande kommt, sie andererseits aber nur dann eingeführt werden, wenn ein breiter Konsens insbesondere bei Lehrkräften und Eltern hergestellt wurde.

Zugespitzt formuliert, könnte man *Comprehensive School Reforms* als Reformen „mittlerer Reichweite“ bezeichnen. Sie gehen deutlich über eine auf Schulfächer oder auf bestimmte Schülergruppen gerichtete Reformidee hinaus, sind aber dennoch klar auf die Einzelschule begrenzt und daher nicht, wie z. B. eine landesweite Lehrplanrevision, von vornherein auf eine Breitenwirkung angelegt.

Betrachtet man vor dem Hintergrund der US-amerikanischen Entwicklungen die Situation in Deutschland, so lassen sich, wenn auch abgeschwächt, einige Parallelen erkennen. Die Gesamtstrategie der Kultusministerkonferenz (KMK) zum Bildungsmonitoring aus dem Jahre 2006 „dient zum einen der systematischen Beschaffung von Informationen über das Bildungssystem. Zum anderen verknüpft sie diese Erkenntnisse mit Maßnahmen zur Unterrichts- und Qualitätsentwicklung“ (KMK, 2006, S. 5). Während der Aspekt der systematischen Beschaffung von Informationen mittels internationaler wie nationaler Studien im Detail dargelegt wird, bleibt der zweite Aspekt eher unverbindlich. Denn die Verknüpfung der Vielzahl von Daten aus den flächendeckenden Vergleichsarbeiten mit Maßnahmen der Unterrichts- und Schulentwicklung wird den Schulen überlassen, welche „eine Fülle von Informationen [erhalten], die sie für ihre pädagogische Arbeit unmittelbar nutzen können“ (KMK, 2006, S. 22). Es ist politisch durchaus nachvollziehbar, dass die KMK ihren Mitgliedern keine Reformstrategien vorschreiben will. Doch in dieser Entkoppelung der Verantwortung für Datenproduktion und Datennutzung liegt auch die Gefahr, dass in der länderspezifischen Umsetzung ein Verwaltungshandeln legitimiert wird, das sich auf teilweise unreflektierte Reformhoffnungen stützt.

Hier mag man einwenden, dass sich in Deutschland, anders als in den USA, das Modell der datengestützten Schulinspektion als ergänzende externe Evaluationsform durchgesetzt hat. Inspektorate werden von den Schulen überwiegend als potenzieller Support wahrgenommen, wenn auch das Feedback der Inspektion zumeist nicht unmittelbar für Zwecke der Schul- und Unterrichtsentwicklung genutzt wird (vgl. Gärtner,

Wurster & Pant, 2013). Nach den oben beschriebenen verzögerten Wirkzeiten (Borman et al., 2003) kann sich erst langfristig erweisen, ob die überwiegend im innerschulischen Konsensmodell entwickelten Änderungsansätze weitreichend genug sind, um substantielle Zuwächse in den Testleistungen der Schüler *auf Landes- und Bundesebene* hervorzubringen. Denn hierfür fehlt eine entscheidende Komponente, die man als den „blinden Fleck“ der Neuen Steuerung im Bildungswesen bezeichnen kann, nämlich die Erhöhung der Reichweite von erfolgreichen Reformen (*Scaling-up*; vgl. McDonald, Keesler, Kauffman & Schneider, 2006). Das Schulinspektionsmodell ist primär auf die Einzelschule bezogen. Es gibt keine administrative Infrastruktur, die es ermöglichen würde, systematisch aus den Einzelerfahrungen zu lernen mit dem Ziel, erfolgreiche Konzepte mit Breitenwirkung zu diffundieren. Das gilt auch dann, wenn man den Professionalisierungsaspekt betont und Schulleitungen als Expertinnen und Experten auffasst. Es fehlen Kommunikationsmechanismen, die es einzelnen oder Gruppen von Schulleitungen erlauben, ihre Erfahrungen in einer Weise einzubringen, dass Schulentwicklungsmodelle auf Landes- oder Bundesebene weiterentwickelt werden. Gleiches gilt für die systematische Dokumentation und Auswertung der Erfahrungen von Lehrkräften bei der Implementation innovativer Unterrichtsmodelle. Weder das alte noch das neue Steuerungsmodell lassen es als wahrscheinlich erscheinen, dass solche Synergieeffekte entstehen. Der Autonomiegedanke der Schule steht solchen Bemühungen tendenziell ebenso entgegen wie eine nach dem bürokratischen Modell arbeitende Administration, die sich auf das Monitoring von Leistungsergebnissen beschränkt.

4. (Aus-)Wege aus einer indifferenten Steuerungspraxis: mögliche Ansatzpunkte für die Bildungsforschung in Deutschland

Die derzeitige Steuerungspraxis im deutschen Bildungswesen erscheint wie „eingeklemmt“ zwischen wirkmächtig gebliebenen bürokratischen Leitvorstellungen und agenturtheoretisch bzw. professionstheoretisch beschreibbaren Antagonismen.

Die schnelle Einführung und rechtliche Verbindlichkeit von Bildungsstandards in der Folge des ‚PISA-Schocks‘ bedingte, dass Implementationsprozesse innerhalb kürzester Zeit und simultan auf mehreren Systemebenen initiiert werden mussten (zentrale und dezentrale Schulaufsicht, Schulleitung, Fachgruppen, Lehrkräfte, Lehreraus- und -fortbildung). Die strukturelle Folge dieser „systemweiten Hast“ ist bis heute eine Asynchronizität und mangelnde Kohärenz zwischen dem Evaluations- und Rechenschaftslegungssystem einerseits und Entwicklungs- und Unterstützungsangeboten für schulische Akteure andererseits, also zwischen den *Pressure-* und *Support-*Elementen. Die Einführung der standard- und ergebnisbezogenen neuen Steuerungsinstrumente, wie Vergleichsarbeiten, zentrale Examina, Schulinspektion oder Verpflichtung zur internen Evaluation, erfolgte *top-down*, implizierte aber gleichzeitig einige notwendige „Nebenbedingungen“ für deren gelingende Implementation: So wurde erwartet, dass sich die „neue“ Unterrichtswirklichkeit quasi durch autopoietisches Engagement der schulischen Akteure *bottom-up* entwickelt, also aus der Profession heraus. Traditionelle,

stofforientierte Unterrichtsskripte sollten in Unterrichtsformen transformiert werden, die kognitiv aktivierende, kompetenzorientierte und kumulativ angelegte Lernprozesse ermöglichen. Die Wirkhoffnungen für eine leistungsrelevante Qualitätsverbesserung gründeten in Deutschland demnach auf dem Paradoxon einer *top-down* initiierten, bürokratisch überwachten *Bottom-up*-, „Strategie“. Die Erwartung, dass Schulen selbsttätig und flächendeckend datengestützte Entwicklungskreisläufe implementieren (vgl. KMK, 2010), findet bisher zumindest keine empirische Unterstützung (Maier, 2010; Maier, Bohl, Kleinknecht & Metz, 2012; Richter, Böhme, Becker, Pant & Stanat, im Druck). Selbst wenn schulische Akteure bereit sind, bei den neuen Steuerungsinstrumenten stärker den Entwicklungsaspekt als den Kontrollaspekt wahrzunehmen, treffen sie derzeit kaum auf kontingente Unterstützungsangebote. Ein handlungspraktischer „Mehrwert“ der ressourcenintensiven Auseinandersetzung mit den Leistungsrückmeldungen aus Schulinspektion, interner Evaluation und externen Tests kann so nicht verbucht werden (vgl. Cobb & Jackson, 2012).

Indifferente oder mehrdeutige Funktionszuschreibungen und Konzeptdefinitionen für neue Steuerungsinstrumente stellen jedoch nicht nur für Lehrkräfte und Schulleitungen ein Problem dar. Dass Schulaufsicht, Schulinspektion, Fortbildner, Studienseminarleiter, Schulleitungen und Lehrkräfte ein halbwegs homogenes Verständnis zentraler Konzepte wie Kompetenzorientierung, schulische Autonomie, Chancengerechtigkeit, Evaluation oder Evidenz entwickelt haben, kann nahezu ausgeschlossen werden. Auf allen beteiligten Steuerungsebenen ist es angesichts der beschriebenen Indifferenz und Mehrdeutigkeit höchst wahrscheinlich zu Prozessen der adaptiven Rekontextualisierung (Fend, 2011) bzw. des „ko-konstruktiven“ *Sense-Making* vor Ort gekommen (vgl. entsprechende Ansätze bei Coburn, 2004; Honig & Rainey, 2012; Mehan, Hubbard & Datnow, 2010; Mintrop, 2012; Sandholtz, 2012). Wie Wacker, Maier und Wissinger (2012, S. 312) ausführen, sind „[d]ie Akteurskonstellationen auf und zwischen den Ebenen, die Wirkrichtungen und die verschiedenen Kopplungen (...) bislang weder hinreichend beschrieben noch in Ansätzen erforscht“.

Für die Bildungsforschung sowohl im Bereich der Steuerungs- als auch der Schulleistungseffektivitäts- und der Schulentwicklungsforschung scheint daher die Diskussion einer Akzentverschiebung sinnvoll und notwendig. Anstelle einer fortgesetzten Fokussierung auf *Output*-Aspekte müsste der theoretischen und empirischen Modellierung von – ebenenübergreifend moderierenden und mediierenden, teilweise reziproken – Prozessen mehr Aufmerksamkeit zukommen. Konzepte wie das der „Pfadabhängigkeit“ eignen sich beispielsweise dafür, systemische „Beharrungstendenzen“ als Prozesse misslungener Handlungskoordination in Mehrebenensystemen zu beschreiben (vgl. Grunow, Pamme, Köhling, Lanfer & Wißing, 2011; Kirchner, 2008; Mason, 2008). Dabei dürfte allerdings die „Endstrecke“ kausaler Wirkungsketten bzw. -konfigurationen, d. h. das Wirksamwerden von Steuerungsimpulsen auf (langfristige) Leistungseffekte, nicht außen vor gelassen werden (vgl. hierzu auch die Schulpanel-Studien von Bischof, Hochweber, Hartig & Klieme, 2013). Besondere Forschungslücken betreffen neben Akteursgruppen in Bildungsadministration (z. B. Schulaufsicht) und Unterstützungssystem (z. B. Coachinganbieter) auch bisher wenig beachtete Akteure wie institutionalisierte

Elternvertretungen (z. B. Beiräte), Bildungsjournalismus, Lehrmittel-Verlage oder zivilgesellschaftliche Akteure im Bildungssektor (Stiftungen).

Die aufgeführte und bei Weitem unvollständige Liste potenzieller Forschungsgegenstände und Fragestellungen legt außerdem nahe, dass die „Handlungskoordination“ rekonstruktiver und deduktiv-quantifizierender Forschungsparadigmen optimiert werden sollte.

Literatur

- Altrichter, H., & Kanape-Willingshofer, A. (2012). Bildungsstandards und externe Überprüfung von Schülerkompetenzen: Mögliche Beiträge externer Messungen zur Erreichung der Qualitätsziele der Schule. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012, Bd. 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 355–394). Graz: Leykam.
- Altrichter, H., & Maag Merki, K. (2010). Steuerung der Entwicklung des Schulwesens. In Dies. (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 15–39). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37, 65–75.
- Amrein-Beardsley, A., Berliner, D. C., & Rideau, S. (2010). Cheating in the first, second, and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives*, 18, 1–31.
- Belfield, C. R., & Levin, H. N. (2009). Market reforms in education. In G. Sykes, B. Schneider, D. N. Plank & T. G. Ford (Hrsg.), *Handbook of Education Policy Research* (S. 513–527). New York: Routledge.
- Bischof, L. M., Hochweber, J., Hartig, J., & Klieme, E. (2013). Schulentwicklung im Verlauf eines Jahrzehnts. Erste Ergebnisse des PISA-Schulpanels. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (S. 172–199). Weinheim/Basel: Beltz.
- Bishop, J. H. (1997). The Effect of national standards and curriculum-based exams on achievement. *American Economic Review*, 87, 260–264.
- Bishop, J. H. (1999). Are national exit examinations important for educational efficiency? *Swedish Economic Policy Review*, 6, 349–398.
- Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73, 125–230.
- Brauckmann, S. (2012). Schulleitungshandeln zwischen deconcentration, devolution und delegation (3D) – empirische Annäherungen aus internationaler Perspektive. *Empirische Pädagogik*, 26, 78–102.
- Brückner, Y., & Tarazona, M. (2010). Finanzierungsformen, Zielvereinbarung, New Public Management, Globalbudgets. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 81–109). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Cabinet Office (1999). *Modernising government*, Cm 4310. London: The Stationery Office.
- Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective. *Journal of the Learning Sciences*, 21, 487–521.
- Coburn, C. E. (2004). Beyond decoupling. Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77, 211–244.
- Ebers, M., & Gotsch, W. (2006). Institutionenökonomische Theorien der Organisation. In A. Kieser & M. Ebers (Hrsg.), *Organisationstheorien* (S. 247–308). Stuttgart: Kohlhammer.

- Fend, H. (2011). Die Wirksamkeit der Neuen Steuerung – theoretische und methodische Probleme ihrer Evaluation. *Zeitschrift für Bildungsforschung*, 1, 5–24.
- Freidson, E. (2001). *Professionalism, the third logic: on the practice of knowledge*. Chicago: University of Chicago Press.
- Fuchs, T., & Wößmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics*, 32, 433–464.
- Fullan, M. (2009). Large-scale reform comes of age. *Journal of Educational Change*, 10, 101–113.
- Gärtner, H., Wurster, S., & Pant, H. A. (2013). The effect of school inspections on school improvement. *School Effectiveness and School Improvement*, 24, 1–20.
- Gonand, F., Joumard, I., & Price, R. (2007). *Public spending efficiency: institutional indicators in primary and secondary education* (No. 543). OECD Publishing.
- Grunow, D., Pamme, H., Köhling, K., Lanfer, J., & Wißing, S. (2011). Verbindlichkeit in einer komplexen Umwelt. Die Kommunalisierung sozialer Hilfen als Gegenstand politik- und verwaltungswissenschaftlicher Forschung. In D. Grunow, H. Pamme, K. Köhling, S. Wißing & J. Lanfer, *Vereinbarte Verbindlichkeit im administrativen Mehrebenen-system* (S. 16–22). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hanushek, E. A., Link, S., & Wößmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, 104, 212–232.
- Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24, 297–327.
- Hargreaves, D. H. (1995). Inspection and school improvement. *Cambridge Journal of Education*, 25, 117–125.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Heinrich, M., Altrichter, H., & Soukup-Altrichter, K. (2011). Neue Ungleichheiten durch Schulprofilierung? Autonomie, Wettbewerb und Selektion in profilorientierten Schulentwicklungsprozessen. In F. Dietrich, M. Heinrich & N. Thieme (Hrsg.), *Neue Steuerung – Alte Ungleichheiten. Steuerung und Entwicklung im Bildungssystem* (S. 271–290). Münster: Waxmann.
- Honig, M. I., & Rainey, L. R. (2012). Autonomy and school improvement: What do we know and where do we go from here? *Educational Policy*, 26, 465–495.
- Jürges, H., Schneider, K., & Büchel, F. (2005). The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. *Journal of the European Economic Association*, 3, 1134–1155.
- Kirchner, S. (2008). Pfadabhängigkeit als Mehrebenenphänomen: Grundlagen und Erweiterungen des Pfadansatzes. *Hamburg Review of Social Sciences*, 3, 317–343.
- KMK (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Luchterhand.
- KMK (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Wolters Kluwer.
- Kotthoff, H.-G. (2003). *Bessere Schulen durch Evaluation? Internationale Erfahrungen*. Münster: Waxmann.
- Koretz, D. (2008). Test-based educational accountability. Research evidence and implications. *Zeitschrift für Pädagogik*, 54, 777–790.
- Koretz, D. (2011). Lessons from test-based education reform in the U.S. *Zeitschrift für Erziehungswissenschaft*, 13 (Sonderheft), 9–24.
- Leithwood, K., & Menzies, T. (1998). Forms and effects of school-based management: A review. *Educational Policy*, 12, 325–346.
- Luhmann, N. (1983). *Legitimation durch Verfahren* (6. Aufl.). Frankfurt a. M.: Suhrkamp.

- Maier, U. (2010). Effekte testbasierter Rechenschaftslegung auf Schule und Unterricht. Ist die internationale Befundlage auf Vergleichsarbeiten im deutschsprachigen Raum übertragbar? *Zeitschrift für Pädagogik*, 56, 112–128.
- Maier, U., Bohl, T., Kleinknecht, M., & Metz, K. (2012). Impact of mandatory testing system and school context factors on teachers' acceptance and usage of school performance feedback data. *Journal for Educational Research Online/Journal für Bildungsforschung Online*, 3, 62–93.
- Mason, M. (2008). What is complexity theory and what are its implications for educational change? *Educational Philosophy and Theory*, 40, 35–49.
- McDonald, S.-K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, 35(3), 15–24.
- McDonnell, L. M. (2013). Educational accountability and policy feedback. *Educational Policy*, 27, 170–189.
- Mehan, H., Hubbard, L., & Datnow, A. (2010). A co-construction perspective on organizational change and educational reform. *National Society for the Study of Education*, 109, 98–112.
- Mintrop, H. (2012). Bridging accountability obligations, professional values and (perceived) student needs with integrity. *Journal of Educational Administration*, 50, 695–726.
- Mintrop, H., & Sunderman, G. (2012). Zentrale Steuerung von Schulentwicklung mit Quoten und Sanktionen: Das US-amerikanische Großexperiment in Schulen für benachteiligte Bildungsschichten. *Die Deutsche Schule*, 104, 8–30.
- Pant, H. A. (im Druck). Aufbereitung von Evidenz für bildungspolitische und pädagogische Entscheidungen: Metaanalysen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*.
- Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29(1), 121–129.
- Richter, D., Böhme, K., Becker, M., Pant, H. A., & Stanat, P. (im Druck). Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten. Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen der Schülerinnen und Schüler. *Zeitschrift für Pädagogik*.
- Sandholtz, K. W. (2012). Making standards stick: A theory of coupled vs. decoupled compliance. *Organization Studies*, 33, 655–679.
- Schimank, U. (1994). *Ressourcenverknappung und/oder Leistungsdefizite: Probleme der westdeutschen Hochschulforschung seit Mitte der siebziger Jahre*. Max-Planck-Institut für Gesellschaftsforschung (MPIFG Discussion Paper 94/2).
- Schimank, U. (2014). Governance und Professionalisierung. In K. Maag Merki, R. Langer & H. Altrichter (Hrsg.), *Educational Governance als Forschungsperspektive* (S. 127–150). Wiesbaden: Springer Fachmedien.
- Schlömerkemper, J. (2009). *Qualitätsentwicklung und pädagogische Professionalität – Perspektiven der Entwicklung einer 'autonomen' Lehrerschaft* (Expertise für das Institut für Qualitätsentwicklung). Wiesbaden: Institut für Qualitätsentwicklung.
- Schmid, K., Hafner, H., & Pirolt, R. (2007). *Reform von Schulgovernance-Systemen. Vergleichende Analyse der Reformprozesse in Österreich und bei einigen PISA-Teilnehmerländern* (= ibw-Schriftenreihe Nr. 135). Wien.
- Stanat, P., & Lüdtke, O. (2008). Multilevel issues in international large-scale assessment studies on student performance. In F. J. R. van de Vijver, D. A. van Hemert & Y. H. Poortinga (Hrsg.), *Individuals and cultures in multilevel analysis* (S. 315–344). Hillsdale: Erlbaum.
- Tenorth, H.-E. (1989). Professionstheorie für die Pädagogik. *Zeitschrift für Pädagogik*, 35, 809–824.
- Terhart, E. (1992). Lehrerberuf und Professionalität. In B. Dewe, W. Ferchhoff & F.-O. Radtke (Hrsg.), *Erziehen als Profession* (S. 103–131). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Thiel, F. (2008a). Organisationssoziologische Vorarbeiten zu einer Theorie der Schulentwicklung. *Journal für Schulentwicklung*, 12(2), 31–39.

- Thiel, F. (2008b). Die Organisation der Bildung – eine Zumutung für die Profession? In Y. Ehrenspeck, G. de Haan & F. Thiel (Hrsg.), *Bildung: Angebot oder Zumutung?* (S. 211–228). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Thiel, F., & Thillmann, K. (2012). Interne Evaluation als Instrument der Selbststeuerung. In A. Wacker, U. Maier & J. Wissinger (Hrsg.), *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung – Empirische Befunde und forschungsmethodische Implikationen* (S. 35–55). Wiesbaden: VS Verlag für Sozialwissenschaften.
- U. S. Department of Education, Office of Elementary and Secondary Education (2002). *No Child Left Behind: A desktop reference*. Washington, D. C. <http://www2.ed.gov/admins/lead/account/nclbreference/reference.pdf> [16. 11. 2013].
- van Ackeren, I. (2003). *Evaluation, Rückmeldung und Schulentwicklung: Erfahrungen mit zentralen Tests, Prüfungen und Inspektionen in England, Frankreich und den Niederlanden*. Münster: Waxmann.
- Vincenti, A., & Geiss, M. (2012). Verwaltete Schule: Geschichte und Gegenwart. In M. Geiss & A. Vincenti (Hrsg.), *Verwaltete Schule* (S. 7–15). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wacker, A., Maier, U., & Wissinger, J. (2012). Zu Bedeutung und Kritik der Steuerungsfor- schung – Fazit und Ausblick. In Dies. (Hrsg.), *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung* (S. 303–322). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Waters, M. (1989). Collegiality, bureaucracy, and professionalization. A Weberian analysis. *American Journal of Sociology*, 94, 45–72.
- Wiesenthal, H. (2000). Markt, Organisation und Gemeinschaft als „zweitbeste“ Verfahren sozialer Koordination. In R. Werle & U. Schimank (Hrsg.), *Gesellschaftliche Komplexität und kollektive Handlungsfähigkeit* (S. 44–73). Frankfurt a. M.: Campus.
- Windzio, M., Sackmann, R., & Martens, K. (2005). *Types of governance in education: a quantitative analysis* (TranState Working Papers No. 25). Bremen: Universität Bremen.
- Wong, K. K. (2013). Politics and governance: Evolving systems of school accountability. *Educational Policy*, 27, 410–421.
- Wößmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65, 117–170.
- Wößmann, L., Luedemann, E., Schuetz, G., & West, M. R. (2009). *School accountability, autonomy and choice around the world*. Cheltenham/Northampton: Edward Elgar.

Anschrift der Autorin/der Autoren

Prof. Dr. Felicitas Thiel, Freie Universität Berlin, FB Erziehungswissenschaft und Psychologie, Habelschwerdter Allee 45, 14195 Berlin, Deutschland
E-Mail: felicitas.thiel@fu-berlin.de

Prof. Dr. Kai S. Cortina, University of Michigan, Department of Psychology,
530 Church Street, Ann Arbor, Michigan, 48109–1043, USA
E-Mail: kai.cortina@umich.edu

Prof. Dr. Hans Anand Pant, Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Unter den Linden 6, 10099 Berlin, Deutschland
E-Mail: hansanand.pant@hu-berlin.de